5

The WWI System Architecture for B3G Networks

Edited by Andreas Schieder (Ericsson GmbH, Germany), Elias Tragos (National Technical University of Athens, Greece), Andrej Mihailovic (King's College London, UK), Jukka Salo (Nokia Siemens Networks, Finland) and Jan van der Meer (Ericsson Telecommunicati, The Netherlands)

5.1 Introduction

The Wireless World Research Forum aims to shape the global wireless future, in which the development of a future network structure plays an important role. The purpose of this chapter is to present such a network architecture vision based on the contributions made by and agreements achieved between five large-scale integrated research projects funded within the European Union's 6th Framework Programme. The contributing projects are WINNER [1], Ambient Networks [2], MobiLife [3], SPICE [4] and E2R [5], which gather in the Wireless World Initiative (WWI) [6].

Developing a reference architecture for future networks requires carefully identification of guiding design principles as intrinsic aspects of the reference model. These architectural principles represent invariants, which need to be carefully chosen to provide sufficient structure to implementations of future network architectures. On the other hand, the model needs to provide enough flexibility and freedom to ensure its future applicability and extensibility. In general, we need to acknowledge that the fixed and mobile networking businesses are becoming mature industry segments, which results in specific requirements being put on the architectures of the telecommunication networks deployed. The most important aspects address the efficiency, extendibility, feature richness, openness and usability. These requirement areas move into focus when the technology has matured such that the intended services are provided to the customer in a basically satisfactory manner. This is the case for fixed and mobile networking, which is a technology area that has moved from its revolutionary phase into an evolutionary phase.

Technologies for the Wireless Future – Volume 3 Edited by Klaus David

These high-level considerations on a future network architecture have led to a couple of architectural principles, which are shared by the five projects gathered in WWI. These principles are discussed in the following section.

5.1.1 Key Architectural Principles

A general trend observable in many modern network architectures is a horizontally layered structure adopting the principles proven successful in the design of telecommunication protocols. A layered structure ensures a decoupling of functional areas and allows the reuse of components as well as a shared usage. While a layered structure is generally acknowledged to be an appropriate design choice, the number of layers and their scope is an important decision to be made for the network reference model we are heading for. The WWI has agreed on four distinct functional layers, complemented by a vertical reconfiguration plane as shown in Figure 5.1, which coexists with the four layers and enhances the control and management functionalities. The reference model assumes different radio access technologies to be realized in the lowest layer as well as a network control layer, a service layer and an application layer placed on top of it. Moreover, the architecture enables the dynamic reconfiguration of each layer through the vertical reconfiguration plane spanning all four layers of the reference model. In general, WWI aims at integrating various radio access technologies in a common framework, ensuring the cooperation of diverse access and core networks and hiding the heterogeneity on the network level to services and applications. The focus of the WWI model is restricted to a scenario that assumes the presence of all project functions contributed to these layers. We thus assume the presence of a RAN as specified by the WINNER project, an ambient control space as specified by the Ambient Networks project, a service layer as detailed by MobiLife and SPICE, as well as the end-to-end reconfiguration enablers and functionalities in the system architecture contributed by the E^{2}R project.

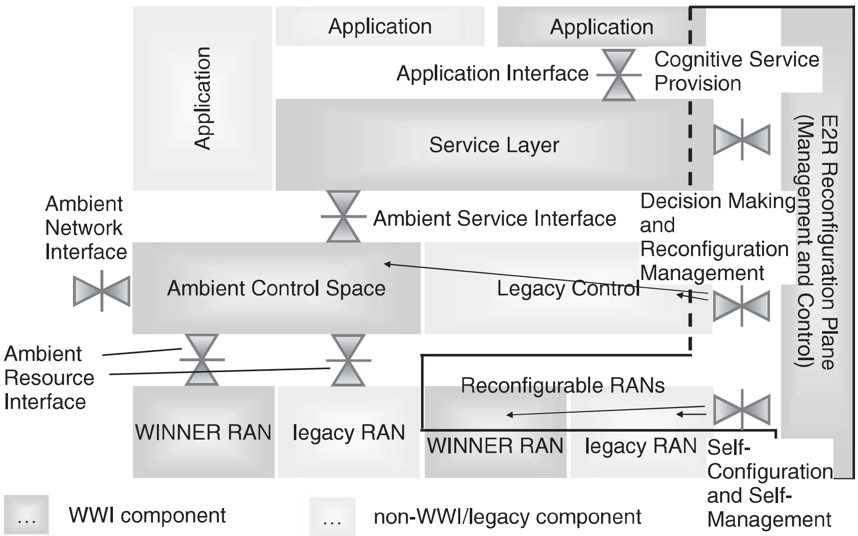

Figure 5.1 Layered structure of the WWI reference network architecture

A specific consideration in the WWI architectural principles is given to the Ambient Networks and E^2R projects in the following text because their scope and relevance extend beyond a single layer in the WWI reference model. Hence, their solutions provide a broad range of interoperability and enhancing features related to the coexistence with and synergy between the WINNER, SPICE and MobiLife projects, as well as with the underlying legacy systems. In the following text we lay out the essential architectural approaches of AN and E^2R in their relations to the WWI reference architecture.

5.1.1.1 Reference Points

Connected to the layered structure of the WWI reference architecture is the presence of some major reference points located between the functional elements.

The ambient network interface (ANI) connects the ambient control spaces (ACS) of different networks (inter-ACS communication), and may also be used to facilitate interaction between functions residing in the same ACS (intra-ACS interaction). The ANI is a reference point either located between different networks or between functions belonging to the same ACS. In the former case, this reference point is made out of a set of interworking interfaces, which will be selectively used based on the purpose of the interworking between the networks. When used between domains the reference point is prescriptive, i.e. it mandates the way in which the objects providing the interfaces are built. When used intradomain, the ANI can be declared to be prescriptive, e.g. for efficiency reasons, however this is irrelevant for the outside of that domain. Otherwise, the ANI is descriptive and thus does not constrain any implementation.

The ambient service interface offers connectivity and control functions for use by upper layer applications within a node operating in a network, or for use by functions residing in the service layer. It allows applications and services to issue requests to the ACS concerning the establishment, maintenance and termination of end-to-end connectivity between functional instances connecting to the ASI. The ASI also includes management capabilities and means to make network context information available to applications and services.

The ambient resource interface is the reference point through which functions of the ACS control and use services of the underlying connectivity. The ARI is located inside a node and also provides an abstraction of the resources in the underlying connectivity infrastructure.

The ARI encapsulates the capabilities of the underlying infrastructure into abstracted information. Because these abstractions are not tied to any specific network technology, the ACS functions are portable between different network types; this decoupling enables a simple migration path from current systems to the WWI system. A main requirement for the ARI is to offer a standard functionality independent of the nature of the underlying technologies.

In the context of the E^2R project, reconfigurability, defined as denoting the capability of a system that can dynamically change its behaviour, tackles the changeable behaviour of wireless networks and associated equipment, specifically in the fields of radio spectrum, radio access technologies, protocol stacks and application services. E^2R features are abstracted in the E^2R reconfiguration plane in Figure 5.1, with management and control functionality involved with all four layers of the WWI reference architecture, with specific focus on RAN reconfiguration aspects, as depicted.

The **Cognitive Service Provisioning Interface** in the overall WWI reference model abstracts the E^2R system architecture functionality of the cognitive service provision as the interface to the service layer. The cognitive service provision module undertakes the necessary content control and service adaptation procedures; finally, after a) sensing the environment and acquiring contextual information and internal status, b) negotiating and deciding the concrete reconfiguration actions, and c) implementing the equipment reconfiguration along with possible adjustment of radio network resources and network elements, the service and content provided to the user may need to be adapted as well.

The **Decision Making and Reconfiguration Management Interface** represents the E^2R interactions with control layers where the actual decision making and reconfiguration management module in E^2R produces and evaluates dynamic policy rules prescribing a set of high-level limitations defining the system behaviour in terms of users and applications requirements, resource availability and business objectives. In addition, it exploits self-knowledge and contextual information in order to formalize the autonomic behaviour of the reconfigurable equipment or network element and produce definitive reconfiguration decisions. Finally, it orchestrates mobility-related signaling, controls reconfiguration sessions, and participates in end-to-end negotiation signaling in use cases involving reconfiguration.

Regarding the interfacing of E^2R reconfiguration plane interactions with RAN layers, the **Self-configuration and Self-management Interface** abstracts the functions of the self-configuration and management (SCM) module, which performs protocol and cross-layer reconfiguration-mode switching, undertakes the self-optimization of local resources, and includes self-healing capabilities. Besides, it orchestrates software upgrade procedures: specifically, it undertakes the actual transfer of downloaded software, the reliable mode switching from many-unicast to multicast and broadcast download modes, as well as local post-download procedures. In addition, it accommodates access and security control mechanisms and generates reconfiguration charging records.

The **Application Interface** makes it possible for applications to access resources on the service layer. Resources are services providing either the required functionality (e.g. sending of instant message and handling related charging) or access to network resources. Both methods are subjected to the access control policies defined for the user of the application interface and the accessed resource. When accessing a resource of the type service, all functions for SPICE service are available as brokering, composition, etc. When the application interface is used on behalf of an end user, where end user-specific data is required on the SPICE platform, the end user must log in to the SPICE platform to handle the appropriate authentication.

5.1.1.2 Main Functional Elements

This overall general WWI reference architecture is complemented by a set of three specific functional architectures addressing key functions foreseen for B3G networks. These functions are:

- *Heterogeneous radio resource management (HRRM)* – refers to the coordination of several parallel-deployed radio access networks to best utilize the resource usage.
- *Mobility management* – refers to the techniques that handle the mobility of the user as he/she is moving between the cells of one network or as he/she is moving between different networks. Traffic balancing and handover are the main techniques (related to the mobility of the users) used in order to efficiently exploit the resources of a network.
- *Context awareness* – a crucial feature for the services, applications and networking functions of the future wireless communications. According to Dey and Abowd [13], context is

any information that can be used to characterize the situation of an entity, where an entity can be a person, place or object that is considered relevant to the interaction between a user and an application, including the user and application themselves.

These functions are described in the following sections.

5.2 Heterogeneous Radio Resource Management (HRRM) in the WWI System Architecture

5.2.1 Introduction and Motivation

Wireless communications utilize radio resources, which typically are limited by physical properties, regulatory requirements (for example spectrum bandwidth and emission powers) and economical constraints. Radio resource management (RRM) concerns control mechanisms to best utilize these resources. For efficiency reasons, RRM is usually fairly tightly coupled to the design of the wireless system it controls. The heterogeneous radio resource management (HRRM) functionality in the WWI architecture targets resource usage efficiency across multiple accesses and faces additional challenges arising from the heterogeneity of the access technologies (in terms of e.g. data rates, coverage, services and existing RRM functionality) and the range of business models to be supported.

5.2.2 Synthesis of a WWI Architecture for HRRM

Realizing HRRM in a network scenario built on the assumption that radio access is provided by various heterogeneous access networks leads to the need to structure the overall network into an access-independent stratum and an access-dependent stratum. This conclusion is supported by all WWI projects, where HRRM represents the fusion of radio resource management functionalities in each involved project (AN, E^2R and WINNER [10], termed as MRRM, JRRM and CoopRRM, respectively). In addition, one can identify a set of functions which need to be present in these two domains to jointly implement the HRRM feature. Our analysis has led to an identification of 12 functions. The purpose of each and whether they are allocated to the access-dependent or access-independent stratum are discussed below.

5.2.2.1 Access-independent Functions

- *Interaccess selection* – this is a generic function that enables the selection of an access resource for scheduling, handover or load balancing purposes.
- *Intersystem access control* – this is a collection of functions for control of the heterogeneous network usage. These functions include load balancing and congestion control amongst networks, intersystem admission control, intersystem handover decision, scheduling and QoS management.
- *Interaccess handover execution* – provides the toolbox to release the current resource and move the call on the target-selected resource within a different access network.
- *Reconfigurability* – coordinates the resulting end-to-end actions in reconfiguration scenarios and, in addition, facilitates network adaptations and spectrum allocation issues.

- *Path selection* – takes into account the end-to-end characteristics of user plane sessions and extends the decision about the access links by functions to select suitable core network paths. The selection takes into account QoS capabilities and security information such as knowledge about the presence of firewalls or private networks.

5.2.2.2 Access-dependent Functions

- *Radio monitoring* – this function enables reporting on the state of current and neighbouring access networks.
- *Access advertisement* – this function enables advertising of the identity and the capabilities of the networks over beacons or other means.
- *Handover execution* – provides the toolbox to release the current resource and move the call to the target-selected resource within the same access network.
- *Access discovery* – this function enables the terminal to discover the available networks.
- *Spectrum management* – this function allows spectrum resource sharing and reassignment between access technologies and operators targeting at a joint optimal resource usage. Benefit for the overall system is the trunking efficiency and load balancing.
- *WINNER deployment mode selection* – as mentioned in [11], WINNER has different operational deployment modes. This function is an internal WINNER function responsible for selecting the target WINNER deployment mode for the users making a handover from legacy networks to WINNER or for new user admission in the WINNER RAN.
- *Reconfigurability* – performs specific reconfiguration actions in the access and terminal parts involving context, QoS, configuration analysis and invoking, profile and policies processing.

All functions mentioned above jointly implement HRRM and ensure the cooperation between heterogeneous radio access networks. In summary, a picture as shown in Figure 5.2 can be constructed to visualize the allocation of the different functions to the two stratums and the radio access network types considered in the WWI context.

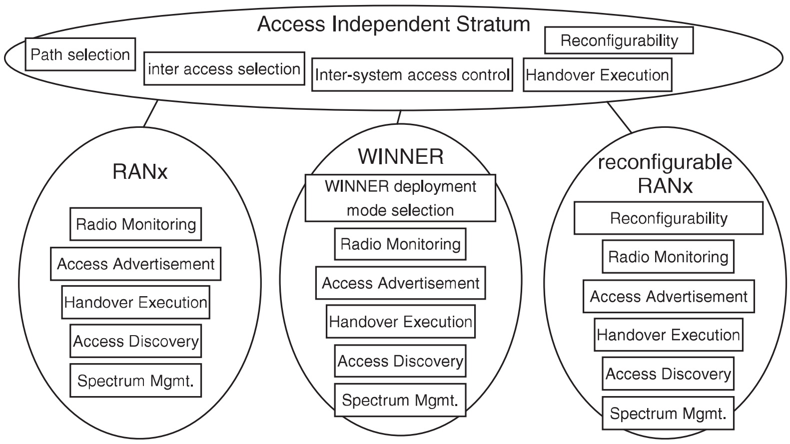

Figure 5.2 First synthesis of a WWI system architecture for the provisioning of multi-radio resource management

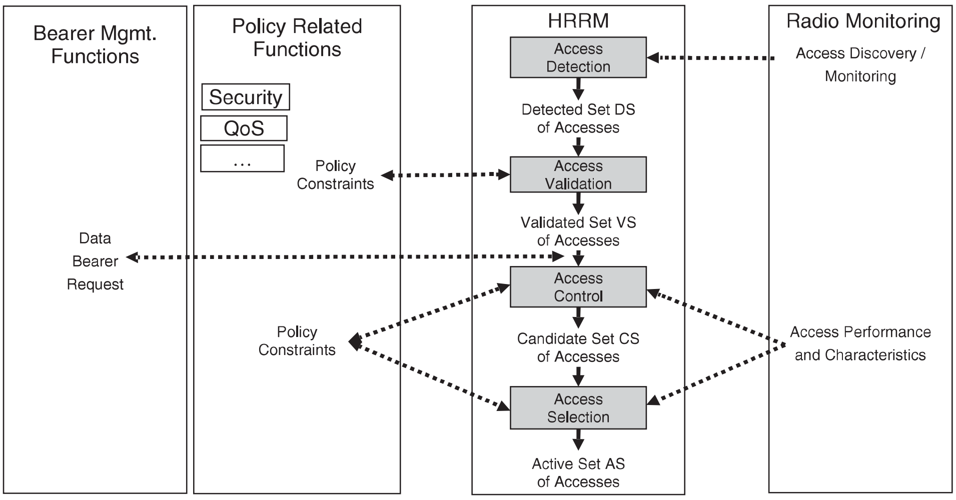

Figure 5.3 Managing different sets in the HRRM

5.2.3 HRRM Decision Making

HRRM is responsible for the selection of an appropriate access network or access link to either satisfy a new connectivity request or to ensure maintenance of an already established connectivity session. The decision on what access to choose is made by executing a sequence of selection steps as depicted in Figure 5.3, which was originally published in [11]. Each of the steps in this sequential process leads to a different set of potential access networks or access links and eventually to the selection of one final access network. The Ambient Networks project has defined four different sets, which are adopted by HRRM. According to [12], these sets are defined as:

- *Detected set (DS)* – this set includes all accesses detected by the mobile terminal either through scanning for available accesses or by listening to a pilot channel broadcasting information about all available accesses. The detected set is maintained by each mobile terminal individually and independent of any current request for connectivity.
- *Validated set (VS)* – the access options maintained in this set are a subset of the accesses present in the DS. Only access networks satisfying certain policy constraints are accepted into the VS. Such policies can include security, QoS requirements and cost restrictions.
- *Candidate set (CS)* – the CS contains a subset of the access options present in the VS. The candidate set is only created after a connectivity request is received and the requirements for the characteristics of the wanted connectivity are known. Access networks capable of satisfying the requirements and possessing sufficient free resources are accepted into the CS. Candidate sets are maintained for each end-to-end connectivity session.
- *Active set (AS)* – the AS is constructed from the CS by applying policy constraints and considering load and performance characteristics of the accesses in the CS. The accesses in the AS are the ones eventually used to implement a particular end-to-end connectivity session.

5.2.4 Use Case Examples

We can distinguish two fundamentally different approaches to implementing HRRM. First, functions in the access-independent and access-dependent stratums may coordinate the usage of different radio access systems to finally ensure that a particular user is served by the best-suited access network. The access networks considered in this cooperative scenario are supposed to deploy different technologies, but are aligned to such an extent that they can interwork and cooperate on the control plane. On the other hand, reconfiguration allows the access networks and/or user terminals to ensure that they always deploy the best-suited technology for the current usage scenario and/or adaptation of networks' operating and/or service delivery conditions.

Variants of the combinations and integration options are explained in the following text by focusing on the explanation of relevant use cases.

The terminal and network cooperation in selecting access networks can be summarized to be provided by the WWI access-independent functions of the HRRM (see Figure 5.4). We distinguish four basic types of access network for the selection process:

- Legacy RANs such as the 3GPP UTRAN or IEEE WLAN and WiMAX networks.
- Reconfigurable legacy RANs, which allow the deployment of different radio access technologies on the same physical hardware and in a shared spectrum.
- Future RANs, represented by the RAN specified by the WINNER project.
- Future reconfigurable RANs.

The reconfiguration of a WINNER RAN is not yet specified, but it is included in the figure as an optional case or as an example of a possible reconfigurable future wireless network. The implementation of the HRRM functions and the interactions between the access-independent functions (HRRM and RCM) are described in the following section by means of a representative set of procedures covering the following use cases:

- HRRM selecting and invocating the WINNER RAN (case 1).
- WINNER RAN and reconfigurable RAN HO request to HRRM for intersystem HO (case 2).
- Spectrum coordination between WINNER RAN and reconfigurable RAN (case 3).
- HRRM/RCM radio reconfiguration of reconfigurable RANs (case 4).
- HRRM/RCM request for access to additional operator resources (case 5).
- Reconfigurable terminal roaming with reconfiguration from one type of network to another (e.g. from Winner RAT to a reconfigurable UTRAN) (case 6).

5.2.4.1 HRRM Selecting and Invocating the WINNER RAN

HRRM is assumed to coordinate the general access network evaluation and selection of radio accesses, while the WINNER RAN also includes radio resource management functions itself to directly coordinate with legacy RANs. When a new radio access is selected, requiring movement of an active session to/from a RAN that is not under direct control of the radio resource management functions in the WINER RAN, the radio access will have to rely on the HRRM control function. The WINNER RAN is treated as any other RAN and included in the process of access selection, as discussed in Section 5.2.3. The access selection procedure

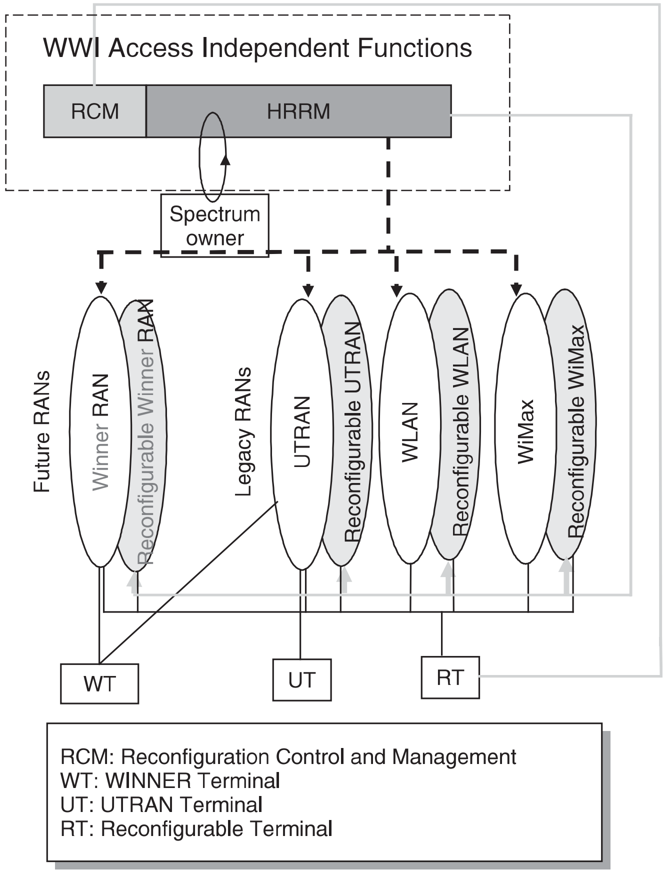

Figure 5.4 Range of HRRM use case deployments

is triggered by a flow setup request that is, for example, received from a client application running on the user equipment. The access selection function in the core network decides on the access network based on the information provided by the mobile terminal in the validated link set (see Figure 5.5).

In the scenario presented here, the WINNER RAN is selected and the access control and access selection functions in the WINNER RAN are triggered to decide upon the best suitable target mode to be applied (see [1] for a discussion of WINNER RAN modes). When the WINNER RAN mode is selected the radio resources are seized, the active link set is updated and the access flow is finally established.

5.2.4.2 WINNER RAN and Reconfigurable RAN HO Request to HRRM for Intersystem HO

The second use case targets the handover of an active session from a WINNER RAN to a second RAN (which is not a WINNER RAN and may be reconfigurable or not). As this is a typical intersystem scenario, the functions in the common core network are mostly concerned

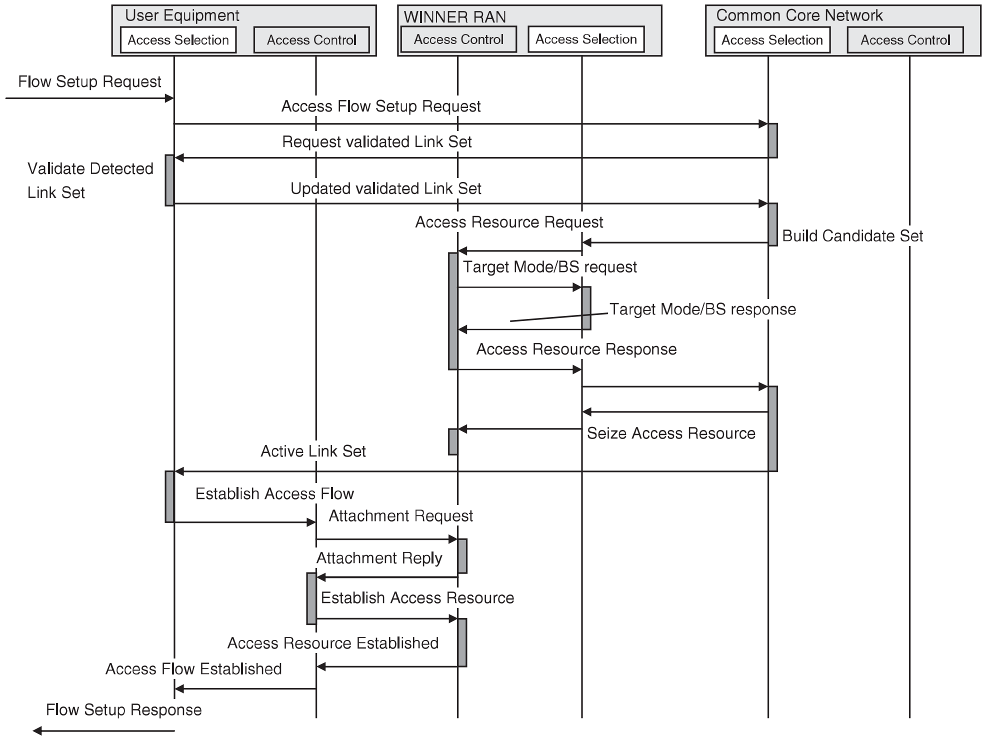

Figure 5.5 Access selection procedure selecting and invocating the WINNER RAN

with the realization of this procedure (see Figure 5.6). The handover scheme employed in the WWI network model relies on the assistance of the mobile terminal to provide information about the access quality (e.g. C/I ratio) to the access selection function in the core network. Based on this information and the core network's knowledge about the end-to-end path, the access selection function in the core network eventually initiates a handover request. The first step in the execution of the handover is the selection of the appropriate handover tool, as we operate in a multitechnology environment and different access networks typically employ different solutions and protocols to execute a handover. The following actual handover execution follows a make-before-break scheme controlled by the handover execution function of the core network.

5.2.4.3 Spectrum Coordination between WINNER RAN and Reconfigurable RAN

HRRM also includes functions to reconfigure radio spectrum usage. Two reconfiguration cases need to be considered: first, spectrum can be assigned to the WINNER RAN to satisfy an increased demand in the WINNER RAN; second, spectrum can be reassigned between reconfigurable legacy RANs to optimize spectrum usage for a particular usage scenario. Both spectrum configuration cases could target the same spectrum and thus lead to contention. There is thus a need for the spectrum owner/regulator to specify resolution rules and perhaps also include a third-party spectrum management provider. The resolution activity is indicated in Figure 5.4, showing HRRM interacting with the spectrum owner (case 3).

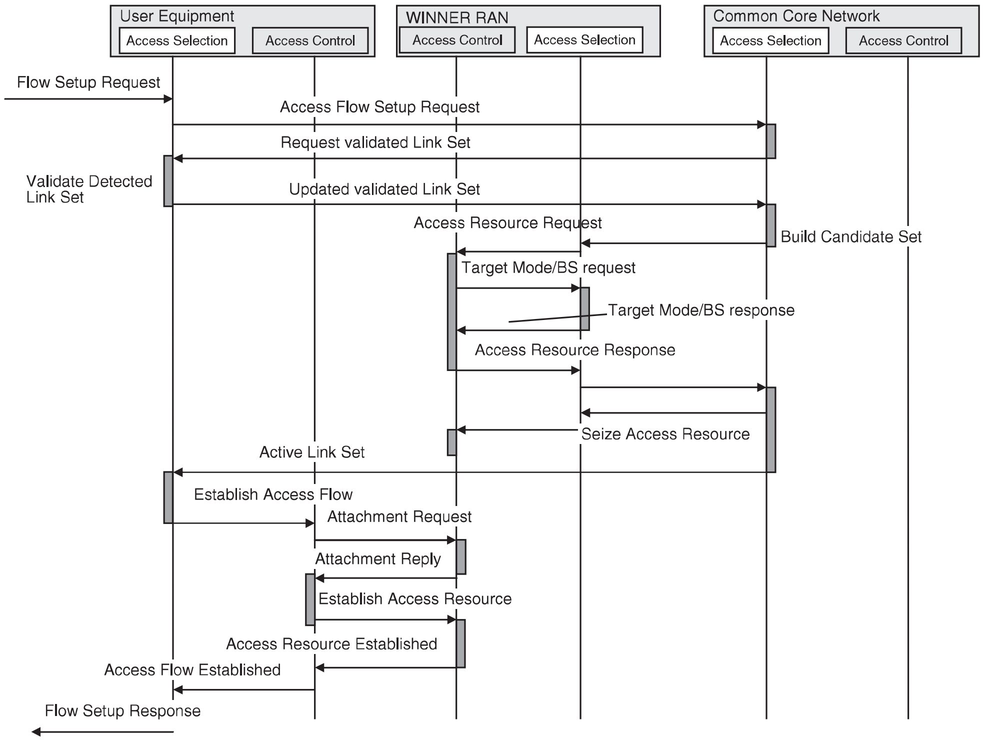

Figure 5.6 Intersystem handover from a WINNER RAN towards a second (reconfigurable) RAN

The spectrum sharing use case can be split into two cases, when the involved RATs in the shared frequency band have equal regulatory status (horizontal sharing), i.e. no system has priority over the other(s) in accessing the spectrum, or when the spectrum sharing is performed with clear established priorities (vertical sharing). In the second case, the primary RAT has a preference in accessing the spectrum and the secondary RAT(s) may only use the spectrum as long as they do not cause harmful interference to the primary.

The horizontal sharing can be split into two cases, namely without or with coordination. In the case where there is no coordination, radio access systems using two or more RATs operate in the same frequency band without the possibility of jointly coordinating their spectrum access. This means that neither system is aware of the location and the current state of the other systems. No signaling is possible between the involved systems, and without the signaling it is generally not possible to prevent interference. However, the only possible way of restricting the effects of mutual interference consists in a high attenuation between the interfering transmitter and the victim receiver and a sufficiently low transmit power. This can be achieved by spatial separation and/or directional antennas. In the case of the horizontal sharing with coordination, the involved radio access systems coordinate their spectrum access based on a set of predefined rules (i.e. spectrum etiquette) that all RATs adhere to. This requires capabilities for signaling or at least detection of the other RATs.

The vertical sharing can also be split into two more cases. In the first case, the WINNER RAT is the primary RAT and the spectrum is dedicated to the WINNER RAT. In this case, no special requirements for the WINNER RAT to coexist with other, secondary RATs exist. The other RATs are allowed to be used only as long as they do not generate interference to any WINNER receiver. The WINNER system can (but is not obliged to) assist the secondary RATs by signaling the free spectrum resources via its broadcast channel while keeping the control of the spectrum utilization. In the second case, the WINNER RAT is the secondary RAT.

The WINNER RAT has to control its emissions (at the base station (BS) and for all UTs) in order to avoid interference to the primary RAT. This requires considerable knowledge of the deployed primary (legacy, non-WINNER) RAT. In this case, some parts of the WINNER system (i.e. short-range cellular, peer-to-peer and feeder links) operate in a frequency band which is assigned to a primary RAT, possibly a legacy RAT. The considered WINNER mode has to operate in such a way that the interference to the primary RAT is prevented. The difference from the horizontal sharing scenarios is that the primary (legacy) RAT has preference in accessing the spectrum and the secondary (WINNER) RAT is not allowed to generate interference, whereas the primary RAT is allowed to interfere with the secondary RAT. The challenge for the secondary RAT, i.e. the WINNER RAT, is to make best use of the 'white spaces' in time, frequency and space, and to transmit in these white spaces only, i.e. without generating harmful interference to the primary RAT. In order to do so, the white spaces have to be identified first and then the transmission parameters have to be adapted accordingly. While the adaptation step is a complex but nevertheless well-defined optimization problem, the identification step is much more difficult to handle and can possibly consist of several components:

1. download from a database (e.g. regulator or another authorized entity)
2. information retrieval from a central radio controller
3. information offered by the primary RAT
4. real-time measurements

Components 3 and 4 could come from HRRM, and components 1 and 2 could be provided in cooperation with E^2R. A combination of these components would give more reliability to the identification process. Naturally, this reliability depends strongly on the properties of the primary RAT. It is much easier to identify the white spaces if the primary system is static in time, frequency and space (i.e. spatial location).

5.2.4.4 HRRM/RCM Radio Reconfiguration of Reconfigurable RANs

When available radio and network resources are limited or specific access is required by user preferences, there are two approaches to increase performance or to meet the user preferences: either evaluate access options, i.e. use HRRM and the reconfiguration control management (RCM) to find additional spectrum (case 4), or launch dynamic roaming agreements through the ambient networks composition process to find additional networks and access options (case 5).

5.2.4.5 Reconfigurable Terminal Roaming with Reconfiguration from One Type of Network to Another (e.g. from Winner RAT to a Reconfigurable UTRAN)

In case of additional radio resources are needed to satisfy the real-time service delivery levels, operators can provide them via the reconfiguration functionality. The actions can include spectrum allocations, dynamic network planning, dynamic upgrading of network components, self healing actions, etc. The use of reconfigurable terminals further provides for the uninhibited roaming of reconfigurable terminals to arbitrary networks (case 6). More specifically, when a terminal wishes to enter into a new network and it lacks a specific set of protocols, it is provided with all the mechanisms to identify which protocols are needed, where to find them, how to download them, how to become reconfigured in real time and in a seamless way, and how to verify the reconfiguration actions.

5.3 Mobility

5.3.1 Introduction and Motivation

Mobility management techniques support user mobility, including the traffic balancing which is essential for a network to use efficiently the resources of the system. In the future, wireless networks will be able to cooperate in order to provide the user maximum QoS, and mobility management is one of the key mechanisms for that. In the WWI system architecture, mobility management will ensure user mobility in heterogeneous access environments.

There are many issues related to mobility that are investigated in the scope of the WWI architecture. The main goal is to define a global WWI functional architecture for mobility management by analyzing and combining the architectures of the different projects, which each deal with mobility from their own perspectives. This will be done by the unification of the mobility requirements and related issues, or to put it another way, by extraction of common and ubiquitous mobility management requirements applicable and present in all three projects, with the ambition of promoting them as essential mobility issues for emerging telecommunication environments.

5.3.2 The Functional Architecture

Mobility management in the WWI concept is related to mobility management between different access networks, since all the projects are working on enabling cooperation of networks, but each from a different perspective. For example, WINNER is working towards making the new network able to cooperate with the legacy networks. Some of the common requirements for mobility management in all the projects are:

- *Handover* – fast and seamless handover execution and need for seamless reconfiguration for the execution of the handover.
- *Triggers* – either for handover and/or for general mobility actions, triggers have been acknowledged in all the projects. The actual triggers vary by and large and can include physical layer-related triggers, context information, user preferences and profiles, location

information, more implicit cases such as adaptation/monitoring of service to system changes, and intelligent triggers.

- *Context and location awareness* – the availability of context and location information to the mobility management processes is another requirement that, if utilized, expedites and improves important performance criteria.
- *Planning* – this relates to the actions that are performed in response to determination that a handover is imminent or in an execution stage. The requirement also demands a fallback to unplanned handovers in case a planned handover fails.
- *Coping with legacy systems and multiple operator deployments* – support for different technologies for vertical handover could relate to seamless execution of reconfigurations for terminals attempting to match to the spectrum and network standard. One extreme case is network reconfiguration due to mass handover requests from terminals with the same radio mode. For example, terminals should be able to become reconfigured when entering into a new access network that belongs to either the same or a different operator. The need for reconfiguration arises from the fact that different networks, especially those belonging to different operators, may use a different set of mobility management, RRM and QoS protocols. Thus, the user equipment needs to be able to be adjustable to the new environment in a seamless way. Obviously, this requires the ability to discover early any incompatibilities between the current configuration and the protocols used by the new network, to locate and download the appropriate set of protocols, and to reconfigure the protocol stack appropriately.
- *Mobility as diverse architectural component* – this includes all of its aspects: session/application, traditional, resource expenditure, load balancing, multi-homing, security, all variants of mobility: endpoints, sessions, flows, interfaces, network/groups, flexible spectrum utilization, group and signaling management [7].

A WWI functional architecture for the handover process part of the mobility management is shown in Figure 5.7. The different functional entities of all the projects related to inter-RAT handover are presented, as well as the connections between these entities. This figure is separated into layers, starting (from bottom to top) at the physical layer and ending at the application and service layers. It provides a good example of the inter-RAT handover process and how the entities of the projects could cooperate in a unified way to handle an Inter-RAT handover. The central entity is the global RRM entity (named HRRM, as it was described in Section 5.2) that is responsible for the RRM decisions, the RAT selection, the load balancing, etc. This figure also shows many triggers for an inter-RAT handover, either from the network layer (from SRRML or SRRMW) or from the service layer (from service or context information). The newest addition is a vertical cross-layer entity related to fuzzy logic, which is described in the next section.

5.3.3 Fuzzy Logic-based Heterogeneous Mobility Management

In this section, the adaptive multicriteria handover algorithm in the WWI concept is presented [7]. The intersystem FLC is responsible for the handover decision. It applies the predefined rules to the current system conditions.

A neural network learns the FLC parameters from the resulting handover quality indicators.

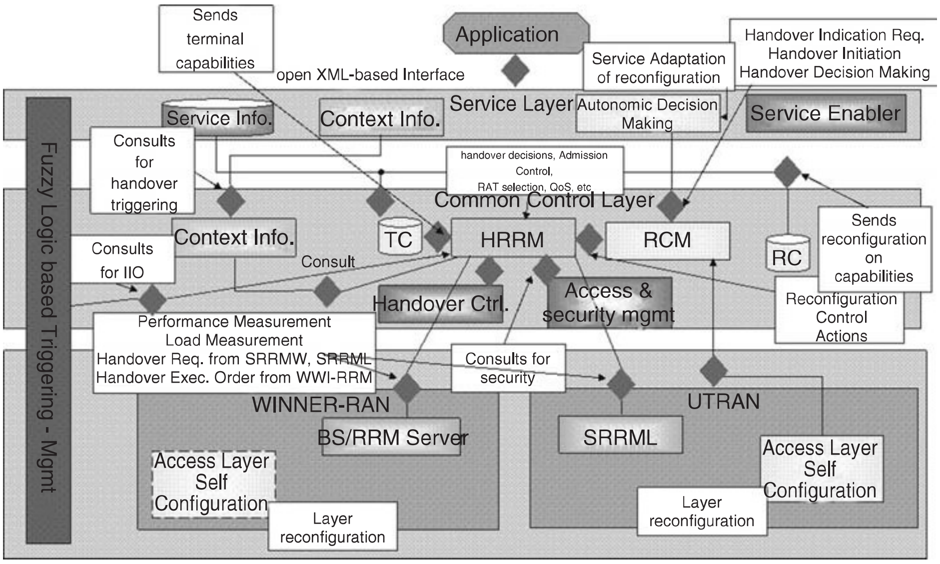

Figure 5.7 Example of functional architecture for the inter-RAT handover process

A preliminary selection of HO targets is performed before the vertical handover procedure. Targets with signal level and load above thresholds are filtered and then the target with the best signal level is chosen. This target preselection reduces the FLC complexity and saves the processing time. The considered handover criteria are signal strength measurements, load information and UT velocity. As they are presented in the Figure 5.8, the triggers for the handover are based on the WWI triggering framework, which considers the mobility triggers from all the projects.

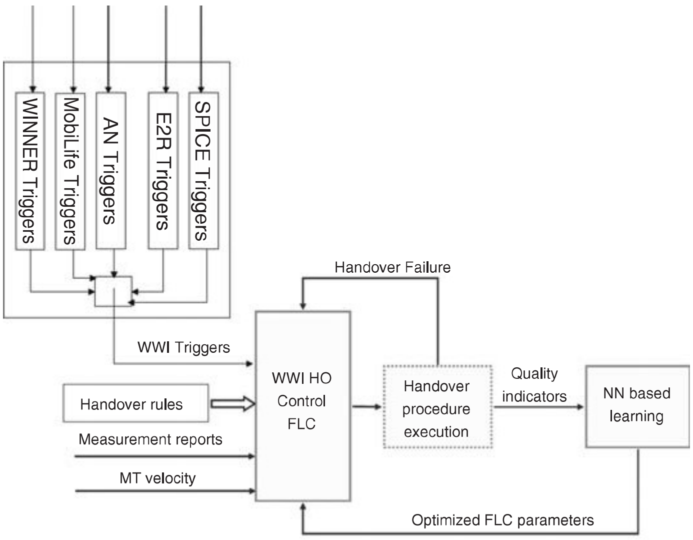

Figure 5.8 Fuzzy multicriteria vertical handover scheme

Neural networks (NN) have been applied successfully to solve complex problems by automatically learning the behaviour of a system and generalizing it to situations not experienced before. A neural network is used in order to determine the optimal FLC parameters whenever the system conditions change. It learns the behaviour of the multisystem network from a database of simulation results.

NN training is very sensitive to learning rate (LR). If it is too high, the training algorithm may be unstable. If it is too low, the convergence will be too slow. Therefore, we use an adaptive LR. The LR adapts to the error evolution during the training process: if the error increases, the LR decreases, and vice versa. Moreover, a momentum is added to back-propagation learning in order to take into account the last weight in the computation of the new weight. The addition of a momentum prevents the back-propagation algorithm from being trapped in a local minimum. The weight correction Δwij applied to the weight of the connection between neuron i and neuron j at iteration n is defined by the following 'generalized delta rule':

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n)$$

n: iteration number
α: momentum constant
η: learning rate parameter
δj: local gradient for neuron j (depends on the nature of neuron j)
yi (n): output of neuron i at iteration n

5.4 Context Provisioning

5.4.1 Introduction and Motivation

Context awareness is a crucial feature for the services, applications and networking functions of the future wireless communications. According to Dey and Abowd [13], context is any information that can be used to characterize the situation of an entity, where an entity can be a person, place or object that is considered relevant to the interaction between a user and an application, including the user and application themselves.

The frameworks for context awareness must provide the context-aware applications with relevant information (where relevancy depends on the user's task) which can be used to adapt the behaviour of these applications to the user's situation. The personalization aspect of context awareness adjusts services to the specific circumstances and needs of a certain person. For example, one probably needs different services while working at the office by the desk and while walking in the street on holiday. A system that knows about the person's situation and activities can be very helpful, not only by providing useful information and services but especially by filtering out those that are unnecessary. Typical context data include spatial information (location, speed, orientation), temporal information (time, duration since an event), environmental information (temperature, light, noise level), user characterization (activity, social surroundings) and resource availability.

Expanding context information a little beyond the user's space makes us consider the network level aspects of context awareness, used for two main purposes. The first is the automatic adaptation of the behaviour of network services to suit the network's situation. The other is the adaptation of end user applications to account not only for user context but also for the relevant information about status and capabilities of surrounding networks within a user's reach.

Reconfiguration of routing media overlays or virtual private networks serving a number of users, dynamic instantiation of functionality to remove bottlenecks or address the need of a community of users, and traffic redirection for load balancing across heterogeneous networks are examples where the dynamic adaptation of networks services is enabled. More intelligent and informed handovers for mobile users and optimized delivery of telecommunication services are other examples showing what network context awareness can enable.

5.4.2 The Usage and Ontologies of Context Information

To manage complexity and account for different requirements at different layers, the context provisioning-related architectures have been studied separately for different communication layers. The scope of the 'context provisioning' aspect of the WWI system architecture is to bring these solutions together, highlighting the interfacing and interworking challenges.

5.4.2.1 Usage of Context Information

Context information can be classified as follows:

- *Human user/group context* – characterizes the situation and circumstances of a user or user group, including their moods, activities/presence, preferences, available devices, environmental information, temporal information, spatial information, etc.
- *Device context* – characterizes the features of the device, including for example screen size and resolution, computational power, battery power and SW capabilities.
- *Network context* – characterizes the resources of the networks and their status, including access types, coverage, bandwidth, supported QoS, supported security functions, etc.
- *Flow context* – characterizes the features of the flow, including the state of the links and nodes that transport the flow, such as latency, jitter, loss, error rate, etc.
- *Service context* – characterizes the service features and current status of running service: required/recommended UIs and multimodality, required execution environment characteristics, service profiles, current mode of activity, current constellation of service components (setup), etc.

The characteristics of the context are:

- Context can grow and become very diverse.
- Context sources may provide huge amounts of data.
- Context sources may be very diverse.
- Context sources have a distributed nature.
- Context information may change frequently.
- Context data may be incomplete, inconsistent and erroneous.
- Relevance of context information can depend on the application and situation at hand.

On the application layer (represented by the MobiLife project), the context information is related to two kinds of entity: individual users and groups of users. Determining the situation of users and groups is necessary to enable situation-aware applications and proactive service provisioning. Context information used in the applications can be related to the location, weather, wellness, presence, preferences, etc. of users and groups.

On the service platform layer (represented by the SPICE project), the context-aware service platform functions – which we refer to as intelligent service enablers – such as context sensing and exchanging, context interpreting, brokering and semantic matching, assist in supporting seamless service and application provisioning for ubiquitous mobile systems. Benefits of the 'intelligence' (the context awareness part) for end users include getting access to personalized and tailored applications in a mobile setting. Within the SPICE project the intelligent service enablers encompass service platform solutions for user profile and contextual information management and anticipatory middleware functionality. Heterogeneous information stemming from context, user profile and service profile sources is processed with advanced reasoning methods targeting at plausible and usable results.

On the network layer (represented by the Ambient Networks project), the networking enablers are needed to achieve a user-centric vision of ambient intelligence. Within these settings the rationale for designing an infrastructure for network context awareness is twofold. On one hand, within very dynamic environments such as ambient networks, there is the need to enhance the operations of other networking functions of the ambient control space to become more capable of reacting to changes in network context in a similar way to how end user applications can be made aware of end user context. On the other hand, the availability of more detailed knowledge about networking context, such as traffic load, QoS, cost, coverage, current performance with regards to specific services (e.g. real-time, file-sharing, etc.), security, etc., is essential in realizing a full contextualization of end user data-communication services and applications.

On the reconfigurability layer (represented by the E^2R project), global knowledge, as well as context information, is needed on connectivity and service reconfiguration. The knowledge about connectivity involves, for example, facts regarding radio propagation characteristics, relationships and theorems defined by the graph theory, and design rules from communication theory. Knowledge related to service provisioning embraces service classes and their mapping to network resources and standards, QoS, trust and reputation rules, strengths of crypto schemes and so on. Consequently, a number of context data, coming from different sources, including not only sensors but also network repositories and devices themselves, are taken into account.

5.4.2.2 Ontologies

Ontologies are formalisms whose purpose is to support humans or machines to share some common knowledge in a structured way. They allow the concepts and terms relevant to a given domain to be identified and defined in an unambiguous way. The ontologies are used to define basic contextual categories and the (logical) relations among them to ensure interoperability in the communication with and between different context providers. The axiomatic descriptions of context elements, such as personal situations (e.g. Working, AtHome, etc.), can directly be used by logical inference engines to realize high-level context reasoning. It is important to note that we do not propose the ontologies described hereafter as the main representation format for all aspects of context modeling, as ontologies are limited to the formulation of qualitative aspects and the available inference engines are generally weak in handling large amounts of data efficiently. Instead, the elements of the XML-based context meta model can be linked to the elements of ontologies to represent qualitative aspects of context information.

Context providers may link context elements representing qualitative data to corresponding elements of the context ontologies. For example, the location provider may allow users to associate location categories to important place (i.e. location clusters) using vocabulary from the ontologies (such as 'home' and 'office'). If a set of qualitative context elements characterizing a user's context, collected from different context providers, is linked to the same set of decidable ontologies, an automatic reasoner can be used to classify the situation.

The goal of the ontologies for supporting context-aware computing applications is that they are simple and easy to understand and use. Currently such ontologies have not been defined which fulfil these basic requirements.

5.4.3 Architecture and Interfaces for Context Awareness

In the pervasive environments, the end-to-end context management (CM) system is needed in order to make context information available in a structured and organized manner on all the layers of the future wireless communication system. Specifically, the context management system needs to collect raw data from available context sources (sensing), process and represent these data into relevant context information (reasoning) and disseminate this information to be used by applications and services (acting). At the same time, such a system has to support the storage and maintenance of context information, as well as their sharing between different system components and applications. The context provisioning functions have to be distributed to different communication layers in a proper way and they have to interact via the commonly defined interfaces as illustrated in Figure 5.9. The system functionality can be divided into the following functional groups:

- *Context source (CS)* – sensors, information services (web).
- *Context provider (CP)* – context data gathering and aggregation, reasoning, distribution, push, subscription, event mechanism, internal data storage for e.g. previously reasoned data,

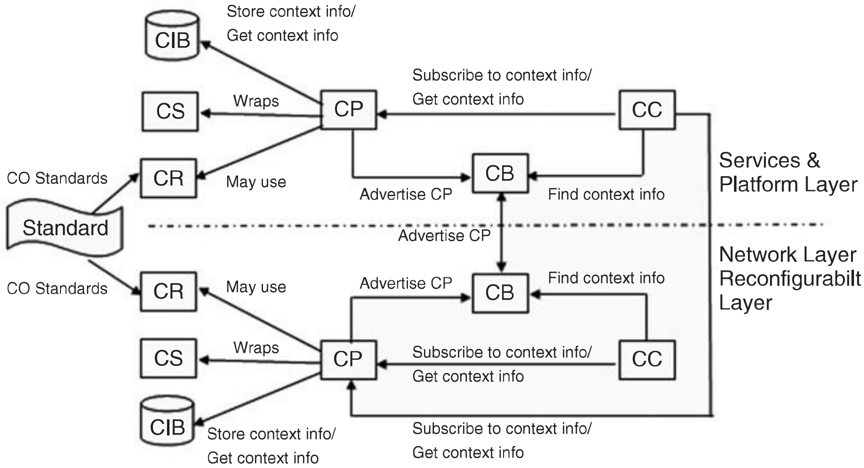

Figure 5.9 System-level architecture for context awareness

context data transformation, security and privacy handling, legal situation handling, context data management.

- *Context client (CC)* – context data query, context consuming.
- *Context broker (CB)* – context provider registration and maintaining updated list of context providers, mapping context requests to context provider addresses, interdomain brokering.
- *Context representation (CR)* – context data model, context reasoning support, ontology.
- *Context information base (CIB)* – storage for inferred context information, storage for basic context information (resources of a network), context history (e.g. sequence of context profiles in time).

5.4.4 Use Cases

The appropriateness of the system-level architecture for the context awareness in services and applications, as illustrated in Figure 5.9, can be demonstrated by some of the use cases. For simplicity, only the following use cases have been selected for presentation in this chapter:

- Register Context Provider.
- Deregister Context Provider.
- Get a Context Provider for an Entity and Parameter.

A key thing for reaching the interlayer interoperability in the presented use cases is that the registration data in the context brokers at the different layers can be maintained synchronized. In this approach the context consumers are able to get the addresses of all context providers from the context broker of their own layer. An alternative for this approach would be that the context brokers know what type of context provider each of them is registering, and when receiving a request from a context consumer they are able to forward that request to the right context broker. In this approach the periodical updates between the context brokers are needed for sharing information on what context items they are authoritative for.

5.4.4.1 Use Case: Register Context Provider

The use case Register Context Provider has been illustrated in Figure 5.10. This use case includes the following main transactions:

- A CP registers to the CB(Network) by sending its advertisement message.
- The CB(Network) makes the registration and forwards the advertisement message to the CB(Service&Platform).
- The acknowledgement is sent to the CP after the successful registration.

The Identification Code (ID) that a CP needs e.g. for deregistering could be ID = URI = CB's URL/ID in CB. Example of ID: http://127.0.0.1.8080/broker/Context Broker/advertisement/3.

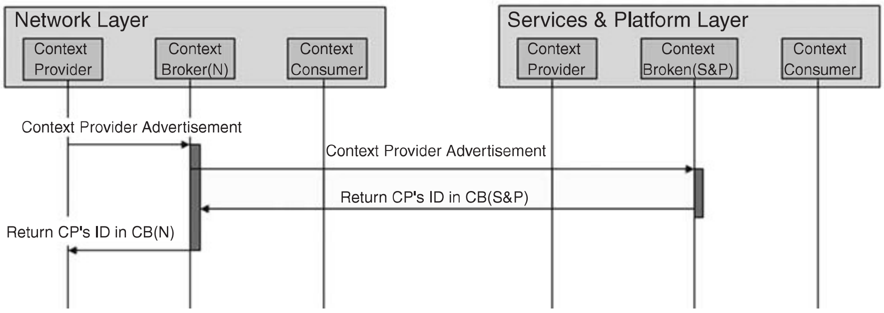

Figure 5.10 Use case: Register Context Provider

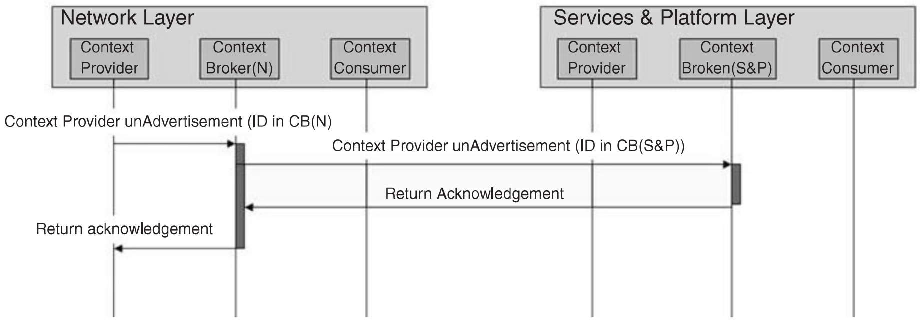

Figure 5.11 Use case: Deregister Context Provider

5.4.4.2 Use Case: Deregister Context Provider

The use case Deregister Context Provider has been illustrated in Figure 5.11. This use case includes the following main transactions:

- A CP deregisters from the CB(Network) by sending its unAdvertisement message with (ID in CB(Network)).
- The CB(Network) makes the deregistration and forwards the unAdvertisement message to the CB(Service&Platform) with (ID in CB(Service&Platform)).
- The acknowledgement is sent to the CP after the successful de-registration.

5.4.4.3 Use Case: Get a Context Provider for an Entity and Parameter

The Use Case Get a Context Provider for an Entity and Parameter has been illustrated in Figure 5.12. This use case includes the following main transactions:

- A CC requests a CP from a CB for wellness information.
- The CB returns the address of a CP (e.g. URL) to the CC.

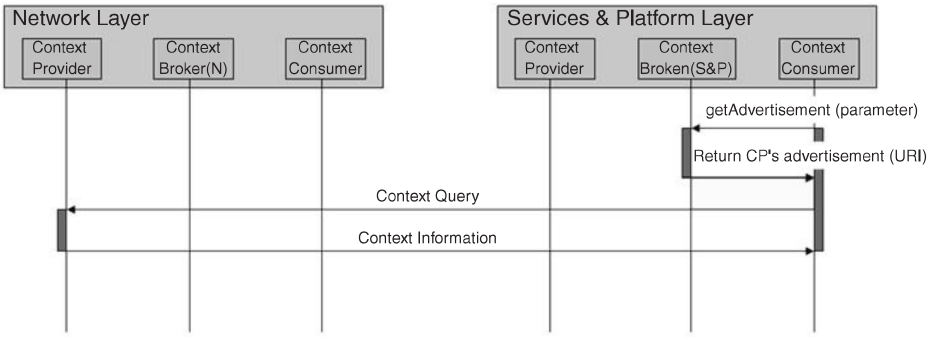

Figure 5.12 Use case: Get a Context Provider for an Entity and Parameter

- The CC requests wellness information directly from the CP.
- The CP returns wellness information to the CC.

5.4.4.4 Challenges

When defining the roles of the communication layers in the end-to-end communication management system and allocating different functions to them, the following challenges have to be paid attention to:

- *Scalability* – the context management mechanisms must be scalable for larger numbers of entities, such as users, operators, devices, different RATs, etc.
- *Data modeling* – the joint representation, storage and exchange of context data has to be standardized. This depends very much on the particular context data, and the respective mechanisms should be obtained from the specific domain; especially in the telecommunication environment, legacy systems already work with their own formats and it is not probable that they change them to conform with one single format. However, a uniform representation of context information must be achieved, and therefore the context management system should take into account mechanisms of mapping to legacy context sources. Since context information is distributed by its own nature, and involves cross-domain interactions, the data model should provide corresponding linking mechanisms to face these issues. Finally, the data model should be easy to work with, understand and represent, as well as allow for reasoning and inferring knowledge.
- *Extensibility* – since the context information can be any information, the mechanisms must be kept extensible, so that the different levels of the system architecture can cooperate and exchange context information.
- *Common ontology* – a common ontology for the different projects must be developed in order to share semantics of contextual information; alternatively, mechanisms for building such a common ontology must be provided. Note that context data model and context ontology are not the same. The former includes the latter, but should also keep such information as the location of context data, its source, its availability, etc.

- *Information base* – the context information base (CIB) is needed to store context and profile information according to a shared ontology. CIB may need to be distributed for performance reasons (to allow both efficient retrieval from context clients and fast updates from context sources).
- *Event mechanism* – the description, creation and distribution of events in the system must be uniform within the system architecture.
- *Privacy and security* – there must be mechanisms in place to protect context information from other operators, users and other entities. These privacy right management solutions and security mechanisms must work across all different architecture levels and involved entities.
- *Legal issues* – different countries have different laws regarding context data and its storage, exchange, etc. It is necessary to define the legal situation for *each* piece of information (e.g. user name, cell ID, preference items, battery level of the phone, URL bookmark, call logs, etc.) and its use with regard to the availability (to the operator, the service developer, the user), the further use and modification, the format (plain, obfuscated, encrypted, etc.), the duration of access, etc. at each stage (sensing, collecting, storing, processing, etc).

5.5 Network Management in the WWI System Architecture

The introduction of the management topic in the context of the WWI system architecture follows the approach to organize, merge and interpret multiple system aspects developed in the involved projects. Applying the management issues in the study of interrelations between the WWI projects, extraction of the ubiquitous conclusions and identification of the differences in approaches, offers specific perspectives on observing and locating the multitude of functionalities involved across the whole end-to-end chain of elements in the emerging telecommunication systems in WWI.

Relevant to this study are the projects AN and E^2R, mainly as their scope implies resolving and developing solutions for the entire end-to-end communications chain and/or impacting all layers of the WWI reference models, either by having explicitly defined management functionalities or by specifying the interfaces for interactions with the layers.

In specific terms, management in the WWI system Architecture refers to the holistic approach in analyzing and converging on the specifics and commonalities of the involved projects and their management/control structures, as well as resolving and understanding the relevant architectural approaches.

5.5.1 Analysis of Main Assessment Criteria

The first observation on the approaches to solving management issues in the overall system architecture of the AN and E^2R projects is based on the identification and analysis of the key assessment criteria typically encountered in general management considerations and development. This is summarized in Table 5.1, with listings of the key management assessment criteria selected and, where applicable, the status of the solutions.

Table 5.1 Status of management assessment criteria in WWI projects

No.	Management Evaluation Criteria	Ambient Networks	E^2R
1	Self-managing networks (self-healing, self-optimizing, self-protecting and self-configuring)	Monitoring and controlling self-configuring algorithms	Overall E^2R system architecture and constituent support mechanisms
2	Self-adaptive applications	Monitoring algorithms, adaptive transcoding of media streams	Multimedia application adaptation
3	Autonomic management of networks, systems and services	Yes	Overall E^2R system architecture and constituent support mechanisms and building blocks
4	Self-repairing distributed systems	No	Self-repairing operations of BS
5	Integrated control and management	Common context and policy planes	Self-ware reconfiguration management plane (S-RMP) concept
6	Distributed, decentralized and scalable management	All management functions are distributed	Mapping of E^2R SA to scalable physical configurations: a) enhancement of 3 GPP SAE networks; b) distribution of FEs to all IP architecture
7	P2P approaches for scalable network management	Overlay management network, P2P resource discovery	UE self-governance: P2P policy information exchange between UE and network policy servers
8	Policy and role-based management	Yes	Dynamic policy generation, policy hierarchies
9	Programmable, active and adaptive management	The monitoring algorithm, policies	OMA DM enablers for: a) setting self-reconfiguration schedule and b) creating DiagMon TRAP.Mo at the device
10	Resilience, dependability and survivability		Disaster recovery scenario
11	'Plug-n-Play' component-based management	Boot-strap function of AN node	Component-based framework for protocol reconfiguration, evolution through the introduction of autonomic protocol component
12	Customer controlled and managed networks	Control of self-x algorithms, scalable monitoring framework.	Optional user interaction during autonomic reconfiguration operations
13	Proactive and reactive management	No	Operation around autonomic control loop for proactive and reactive actions
14	Biologically-inspired management systems and techniques	No	Bio-inspired algorithms for spectrum and radio resource management

In the following, the selected key management criteria are discussed in more detail regarding the essential explanations of the solutions and the fitting approaches to the involved project(s).

1. *Self-managing networks (self-healing, self-optimizing, self-protecting and self-configuring)* – AN: monitoring and controling self-configuring algorithms; E^2R: the E^2R II SA incorporates functionalities for a) self-healing UEs and base stations (through automatic recovery techniques), b) self-optimizing protocols and protocol stacks (through autonomic component replacement and optimization heuristics), c) self-protecting UEs (through secure sandbox mechanisms and secure component replacement), d) self-configuring UEs (through autonomic RAT-switching at the UE and base stations).
2. *Self-adaptive applications* – AN: self-configuring FE provides a point of control for localized optimization algorithms. This FE captures the set commands issued by external algorithms and provides a set of filtering actions to limit or shape the new configuration values. In this value, stability of configuration can be assured; E^2R: E^2R considers the case where real time multimedia application sessions cannot be maintained during the reconfiguration process (or during part of this process) and proposes an automatic set up of new sessions (based on new QoS characteristics) once the new configuration is enforced, for continuing the previous multimedia sessions which have been interrupted.
3. *Autonomic management of networks, systems and services* – E^2R: E^2R II SA incorporates dedicated modules and functional entities for autonomic management of networks/systems (cf. knowledge and context management modules, decision-making and reconfiguration management, self-configuration and self-management building blocks) and services (cf. cognitive service provision building blocks).
4. *Self-repairing distributed systems* – AN: registries management FE provides a highly resilient general distributed registry service for functional entities. An important self-repairing system component is the efficient and self-stabilizing overlay maintenance algorithm of the underlying DHT in registries management FE; E^2R: E^2R introduces mechanisms to assure the connectivity of the base station to the rest of the cellular network, opting for auto-administration of the reconfigured BS through the incorporation of self-repairing actions. This requires discovery of the candidate network point to attach to (BSGW) and the association between BS and the BSGW, as well as transport capacity negotiation between the BS and the controller/gateway.
5. *Integrated control and management* – AN: management in AN is not working in a completely separate management plane but is integrated into the ambient control space (ACS); E^2R: the overall E^2R system architecture is based on the concept of a self-ware reconfiguration management plane (S-RMP), viewed as a unified control and management framework for the coordination of end-to-end interactions between the involved entities, and for enabling the decision-making and enforcement of mechanisms supporting reconfiguration in a dynamic fashion.
6. *Distributed, decentralized and scalable management* – AN: distributed and decentralized management for scalability is one of the key concepts of each component of the integrated AN management layer; E^2R: the E^2R II SA has been mapped to the 3GPP SAE network, which provides scalability through the hierarchical structure of control elements. In addition, mapping to an All-IP network configuration facilitates distribution of functional capabilities via flatter architectures. Mass upgrade mechanisms bear inherent scalability properties.

7. *P2P approaches for scalable network management* – AN: the use of DHTs for AN registries, AN monitoring; E^2R: E^2R foresees the migration from the traditional policy control model between the network policy decision function and the UE policy enforcement function, towards a model whereby the UE and the network operate in P2P fashion.
8. *Policy and role-based management* – AN: the policy management FE provides node- and AN-level policy management services for all functional entities in the ACS. Policies are stored in nodes (node-level) and in AN-wide registries provided by the registries management FE (AN-level); E^2R: E^2R proposes dynamic policy generation, taking into account defined policies and contextual information with the use of ontologies. E^2R foresees prioritization of policies with the use of policy hierarchies, defining certain reconfiguration actions in an autonomic way.
9. *Programmable, active and adaptive management* – AN: the monitoring algorithms adapt to the dynamics of the monitored variables and the network environment; E^2R: E^2R exploits OMA DM enablers for such management features, and specifically the possibility to set a self-reconfiguration schedule in the device to reconfigure itself regularly or according to certain situations, or even create DiagMon TRAP/MO at the device.
10. *Resilience, dependability and survivability* – E^2R: E^2R exploits resilience, dependability and survivability considering disaster recovery scenario. For example, in the case of a natural disaster, one may anticipate that whole areas become inaccessible and some deployed access points may become completely devastated. E^2R proposes mechanisms that will enable remaining access points to be reconfigured in an autonomous and intelligent manner that will also improve the system integrity and maintainability. This way the system can, even at short notice, become once again operational and enhance its availability and reliability by offering vital services to users in a safe manner.
11. *'Plug-n-Play' component-based management* – AN: a collaborative mechanism is defined for wireless base stations. A wireless network can be deployed through a plug-n-play mechanism: base stations sense and discover each other and negotiate the optimal configuration, like channel assignment; E^2R: E^2R proposes a component-based protocol stack approach. The introduced framework aims to cope with the dynamic binding of component services into a fully fledged protocol service and the runtime replacement of protocol functionality. An enhancement of this solution is also proposed with the introduction of autonomic protocol components, reducing the functionality of centralized management and introducing intelligence to the protocol components.
12. *Customer controlled and managed networks* – E^2R: the introduction of E^2R reconfigurability classmarks allows the user to interact with the terminal equipment during different stages of the reconfiguration process.
13. *Proactive and reactive management* – E^2R: the E^2R II system architecture is built around an autonomic control loop that facilitates real-time decision-making dependent on both past observations and online measurements.
14. *Biologically-inspired management systems and techniques* – E^2R: E^2R exploits general heuristics (meta-heuristics) inspired by a model of division of labour used by social insects such as ants and wasps, and incorporates this model in an innovative and efficient distributed spectrum allocation algorithm, called Distributed Agent-Based Variable Threshold OFDMA (DABVT-OFDMA), whose objective is to maximize the system throughput by exploiting multiuser.

5.5.2 Management Planes

Following the above identification of the major management aspects in the involved WWI projects, E^2R and AN, we highlight the specific issue of management plane(s) and management plane overlay(s) as functional placeholders for the collective sets of features and functionalities.

5.5.2.1 E^2R Management Plane Overlay

The E^2R project envisages cognitive networks as the paradigm of heterogeneous next-generation mobile systems that exploit SDR/CR and consist of terminal equipment and network elements with autonomic decision-making and self-management behaviour, while being capable of enriching their knowledge and generating dynamic policy rules based on contextual information. The overall system architecture is based on the concept of a self-ware reconfiguration management plane (S-RMP), viewed as a unified control and management framework for the coordination of end-to-end interactions between the involved entities, and for enabling the decision-making and enforcement of mechanisms supporting reconfiguration in a dynamic fashion.

E^2R specified the E^2R II system architecture for autonomously reconfigurable user equipment and network elements, integrating functional entities for software and cognitive radios while capturing autonomic communication aspects. Such work resulted in the identification of four S-RMP modules (Figure 5.13): the knowledge and context management (KCM) module for knowledge and context retrieval and interpretation; the decision-making and reconfiguration management (DM&RM) module for decision making and orchestration of reconfiguration operations; and the self-configuration and management (SC-M) module for applying self-ware procedures, altogether yielding cognitive service provision (CSP) offering to end users.

5.5.2.2 AN Management Plane Overlay

Amongst the design principles of ambient networks there is a new approach to providing the functionality that has been traditionally created in the management plane in the past. Such

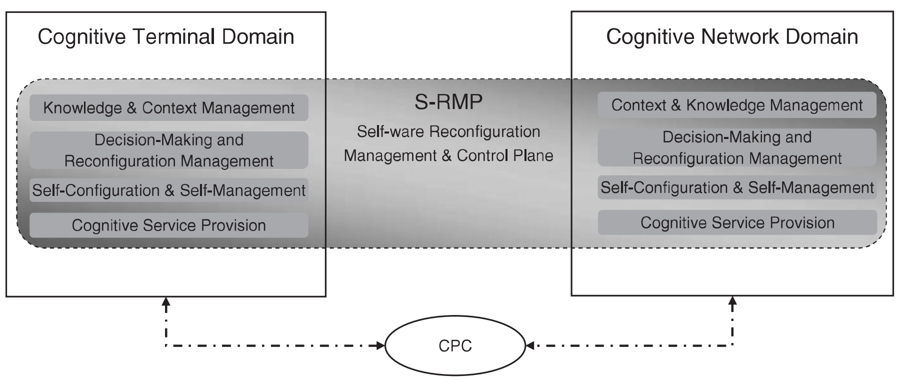

Figure 5.13 E^2R II high-level system architecture for cognitive reconfigurable wireless networks

a novel approach advocates the existence of integrated management solutions that support cooperating networks in highly dynamic scenarios where the need for autonomic behaviour is paramount to reduce complexity and operating costs.

These solutions are to allow some level of network functions self-organization and reconfiguration, based on a tight interaction between a network context management system, a scalable policy framework and an underlying network monitoring system.

Through continuous awareness of underlying network features, resources, faults and performance while delivering various types of applications, ambient networks become more service aware and eventually capable of dynamically reconfiguring the allocation of resources and enabling their efficient usage.

Moreover, distributed monitoring of network performance and the subsequent mapping of network context and network management information to higher-level policies enable the creation of a set of constraints that limit the freedom of operation for the functional entities (FE) composing the ambient control space.

As we anticipated, in order to achieve such objectives, ambient networks depart from the traditional centralized management approach to account for more distributed ways of collecting monitoring information, aggregating it and reacting to it.

This is realized through the concept of the ambient network integrated management node (see Figure 5.14). As the picture shows, the role of management in ambient networks is carried out collectively through a set of functional components as follows. The monitoring FE provides network-wide metrics, representing the state of the network considered as

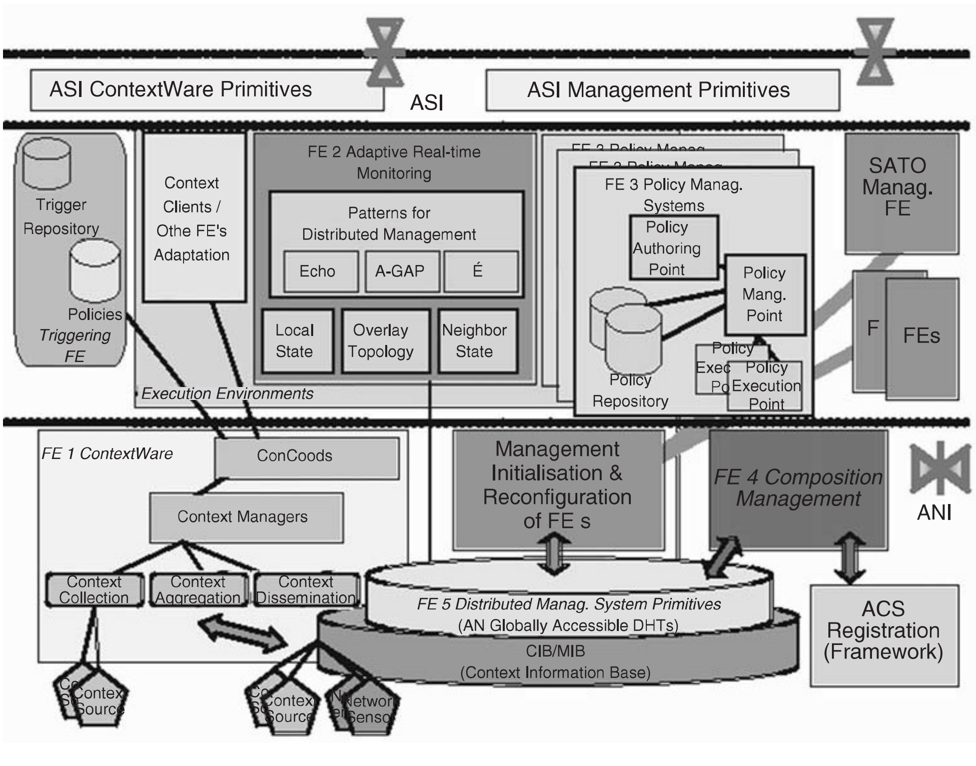

Figure 5.14 AN management node model

a whole. ContextWare FEs confer to ambient networks to become context aware, or to be more precise, aware not only of the services that are being delivered but also of the network's continuous ability to support their requirements. TRG FE and related ANISI (ambient networks information service infrastructure) empower ambient networks with a subscription/notification mechanism enabling clients to quickly react to triggers that cross given thresholds. The self-configuration FE provides general services to other FEs to correctly accomplish their configuration tasks. Its main purpose, currently investigated, is to guarantee stability of a certain configuration over time and to avoid oscillations. The policy FE facilitates the role of authoring and injecting policies which then provide additional constraints to the behaviours of other management components, as well as other FEs in the ambient control space. The composition management FE manages the logical structure of the ACS, as well as FE instances in this structure. Finally, the goal of a common AN distributed registry (registry management FE) is to exploit synergies and provide a unified distributed registry functionality for all functional entities, as a number of functional entities need registries to store and look up various entities.

In order to allow for a high degree of flexibility, it is envisaged that such nodes will be dynamically deployed in the most appropriate locations in a network topology through a 'virtual management backbone' overlay. Each individual management node part of this overlay may not necessarily implement the complete set of management functionality, but only the required subset according to what particular management task needs to be carried out.

5.5.3 Analysis of Management Plane Overlays

The conducted considerations of the management solutions and specifics in AN and E^2R projects reveal the extent and types of difference in approach of the two projects, but also point out the commonalities in addressing the similar management issues, as captured in the management assessment criteria, either by providing specific or, in fact, offering similar solutions.

The starting point in the analysis of the two projects is their guiding general concepts for advancing the operations and orchestrations of telecom systems, AN introducing the concept of ambient connectivity based on the notion of network compositions, E^2R introducing various types of adaptation actions based on the notion of end-to-end reconfigurability. Regarding the implementations of the management planes, AN introduces a distributed approached by integrating the management aspects of cooperating networks and proposes a model based on the AN management nodes which contain the relevant collections of functional entities that govern the management operations and are located and distributed in the network topology via a 'virtual management backbone' overlay. E^2R introduces the dedicated management plane called self-ware reconfiguration management plane (S-RMP), which orchestrates the whole set of end-to-end reconfiguration actions and contains the functional blocks for resolving the decision-making and invocation of functions supporting reconfigurations. The models in which the management tools are realized in the two projects can be associated with the scope and general guiding concepts applied in the projects. As advocated in the projects, these models provide the most efficient way of placing and organizing those functionalities in the system. Hence, if an observation is made on the possible common and integrated WWI management plane structure, the specifics and tools realized in the two projects would have to be initially included in the form they are developed in the projects, particularly as they are the required models for targeting the relevant solutions. Then, in such a coexisting environment, a further study could be conducted on resolving the commonalities and solving the operational conflicts (if applicable). It seems that in such scenarios the current features of the management solutions in the projects could individually resolve or at least accommodate the integration of the two sets of solutions.

Finally, related to the details and options for resolving the management objectives, the identification of the solutions for the management assessment criteria presents the manner in which both projects tackle them. However, it can be highlighted that both projects apply some common management principles related to the awareness and implementations of degrees of autonomics and self-management in the management of system operations. In fact, by observing the specific functionalities in both projects, autonomics and self-management are becoming an inherent aspect of their management functionality.

5.6 Conclusions

This chapter presents a system architecture for B3G networks, which are legacy-respecting networks building on current 3G cellular networks and integrating telecommunication and data networks in the fixed and mobile domain more closely than 3G systems already do. The architecture presented here follows a horizontally-layered design distinguishing basic connectivity, comprising access and core/backbone connectivity, a network control layer, a service layer on which the end-user applications reside, as well as a vertical reconfiguration plane spanning all of the horizontal layers mentioned before. Apart from the presentations of an overall reference model explaining the different layers and the reference points residing between them, four functional areas found to be of key importance to B3G networks are discussed in more detail. These functional areas are: heterogeneous radio resource management, mobility management, context awareness and management. The architecture presented here is the result of a joint activity of five large-scale integrated projects executed in the 6th Framework Program of the European Commission. These projects have gathered under the umbrella of the Wireless World Initiative (WWI) to jointly develop a vision for B3G.

5.7 Acknowledgements

The editors would like to especially thank Imran Ashraf (Alcatel Lucent, UK), Giovanni Bartolomeo (University of Rome 'Tor Vergata', Italy), Zachos Boufidis (University of Athens, Greece), Khadija Daoud (Orange Labs, France), Alex Galis (University College London, UK), Raffaele Giaffreda (BT, UK), David Grandblaise (Motorola, France), Oliver Holland (King's College London, UK), Alexandros Kaloxylos (University of Athens, Greece), Herma van Kranenburg (Telematica Instituut, Netherlands), Jijun Luo (Siemens, Germany), Juha Mikola (Nokia, Finland), Markus Muck (Motorola, France), Eleni Patouni (University of Athens, Greece), Mikael Prytz (Ericsson, Sweden), Abed Samhat (Orange Labs, France), Ove Strandberg (Nokia Siemens Networks, Finland), Anthony Tarlano (DOCOMO Eurolabs, Germany) and Dorota Witaszek (Fraunhofer Institute FOKUS, Germany) for the contributions to this work.

References

[1] http://www.ist-winner.org/.

[2] http://www.ambient-networks.org/.

[3] http://www.ist-mobilife.org/.

[4] http://www.ist-spice.org/.

[5] http://e2r2.motlabs.com/.

[6] http://www.wireless-world-initiative.org/.

[7] WWI Cross Issue System Architecture Mobility Management Final Report.

[8] D. Bourse, M. Muck, O. Simon, N. Alonistoti, K. Moessner, E. Nicollet et al., "End-to-End Reconfigurability (E^2R II): Management and Control of Adaptive Communication Systems", *IST Mobile Wireless and Communication Summit Conference*, Greece, Mykonos, June 2006.

[9] J. Zoric, E. Postmann, J. van der Meer, J. Hendriks, W. Kellerer, T. Srisakul and J. Rovira, "Service and Information Roaming – Architecture and Design Aspects", WWRF 17, Heidelberg, November 2006.

[10] A. Schieder, E. Tragos, A. Mihailovic, A. Kaloxylos, M. Prytz, O. Strandberg et al., "Mobility Management and Radio Resource Management in the WWI System Architecture", WWRF 17, Heidelberg, November 2006.

[11] J. Lundsjö, R. Aguero, E. Alexndri, F. Berggren, C. Cedervall, K. Dimou et al., "A Multi-radio Access Architecture for Ambient Networking", *14th IST Mobile and Wireless Communications Summit*, Dresden, Germany, 19–23 June 2005, available at http://www.ambient-networks.org/phase1web/main/publications_incl_Y2.html.

[12] G.P. Koudouridis, R. Aguero, K. Daoud, J. Gebert, M. Prytz, T. Rinta-aho et al., "Access Flow Based Multi-radio Access Connectivity", *PIMRC 07*, Athens, Greece, September 2007.

[13] A.K. Dey, G.D. Abowd and D. Salber, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-aware Applications", *Human-Computer Interaction*, **16**, pp. 97–166, 2001.